

Channel Capacity of an Array System for Gaussian Channels With Applications to Combining and Noise Cancellation

K.-M. Cheung and V. Vilnrotter
Communications Systems and Research Section

A closed-form expression for the capacity of an array of correlated Gaussian channels is derived. It is shown that when signal and noise are independent, the array of observables can be replaced with a single observable without diminishing the capacity of the array channel. Examples are provided to illustrate the dependence of channel capacity on noise correlation for two- and three-channel arrays.

I. Introduction

In this article, we formulate the framework to evaluate the channel capacity of an array system. We define the channel capacity of an array channel as the maximum of the mutual information between a single input source and an array of n output observables over all distributions on the input that satisfy certain constraints (e.g., power, bandwidth, discrete or continuous, etc.). We derive a closed-form general formula to the channel capacity of an array of n Gaussian channels. This formula applies to correlated and uncorrelated noise, as long as the Gaussian assumption holds and the second-order statistics (covariance) of the signal and noise sources can be characterized. This formula allows one to find the channel capacity of various array constellations and orientations in the presence of correlated and uncorrelated noise. Some of the interesting results that we observed are as follows:

- (1) When the noise sources are uncorrelated, the array channel capacity is equivalent to the channel capacity of a single Gaussian channel with a signal-to-noise ratio (SNR) equal to the sum of the SNRs of the individual channels.
- (2) The array channel capacity is lower bounded by the channel capacity of the channel with the highest SNR.

II. Problem Formulation

We consider transmitting a single complex source through n channels, as illustrated in Fig. 1. Let $\sigma_{X,ij} \triangleq EX_i X_j^*$, $\sigma_{Z,ij} \triangleq EZ_i Z_j^*$, and $\sigma_{Y,ij} \triangleq EY_i Y_j^*$ for $1 \leq i, j \leq n$. Notice that $\sigma_{X,1}^2, \sigma_{X,2}^2, \dots, \sigma_{X,n}^2$ are the signal powers as seen by the receivers; $\sigma_{Z,1}^2, \sigma_{Z,2}^2, \dots, \sigma_{Z,n}^2$ are the noise variances; and $\sigma_{Y,1}^2, \sigma_{Y,2}^2, \dots, \sigma_{Y,n}^2$ are the channel output variances. The array channel capacity is given by

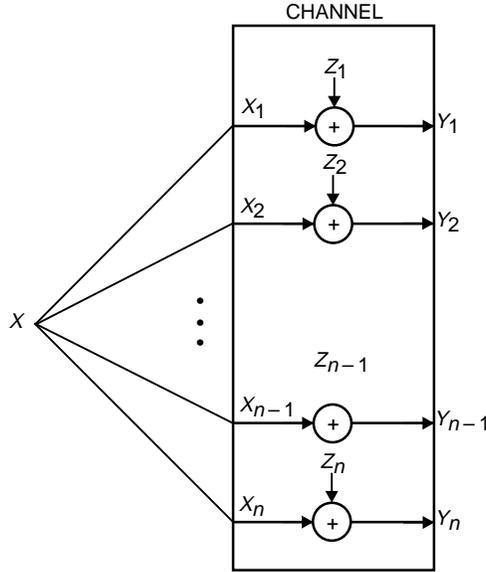


Fig. 1. Array channel model.

$$\begin{aligned}
C_{array} &\triangleq \max_{p(x)} I(Y_1, Y_2, \dots, Y_n; X) \\
&= \max_{p(x)} \{H(Y_1, Y_2, \dots, Y_n) - H(Y_1, Y_2, \dots, Y_n|X)\} \\
&= \max_{p(x)} \{H(Y_1, Y_2, \dots, Y_n) - H(Z_1, Z_2, \dots, Z_n)\} \tag{1}
\end{aligned}$$

where we further assume a power constraint on X , and where the X_i are obtained from X by a deterministic operation on each i . The interpretation of this model is that signals in the various channels can have different magnitudes and phases, but that the differential delays between the waveforms are negligible, having been removed by delay compensation. Thus, the signal path differences between the various channels either are negligibly small, as in the case of array-feed or phased-array applications, or have been properly equalized, as in the case of antenna arrays.

III. Capacity of an Array of Gaussian Channels

In the following derivation, we will use some results from Cover and Thomas [1]. Since $\underline{Z}^T = (Z_1, Z_2, \dots, Z_n)$ is a complex Gaussian random vector, $H(\underline{Z})$ is given by [1, Theorem 9.4.1]

$$H(\underline{Z}) = \frac{1}{2} \log_2(2\pi e)^n |\Theta_Z| \quad \text{bits/channel use} \tag{2}$$

where Θ_Z is the covariance matrix of \underline{Z} , and $|\Theta_Z|$ is its determinant. From Theorem 9.6.5 of [1] and under the assumption that the input source is power constrained to P , the input source X that maximizes $H(Y_1, Y_2, \dots, Y_n)$ has a Gaussian distribution. The maximum mutual information and, therefore, the array channel capacity are achieved with a Gaussian source, and the output observables Y_1, Y_2, \dots, Y_n are correlated complex Gaussian variables. Using Theorem 9.4.1 of [1], $H(Y_1, Y_2, \dots, Y_n)$ is given by

$$H(Y_1, Y_2, \dots, Y_n) = \frac{1}{2} \log_2(2\pi e)^n |\Theta_Y| \quad \text{bits/channel use} \quad (3)$$

where Θ_Y is the covariance matrix of Y_1, Y_2, \dots, Y_n . The array channel capacity is

$$C_{array} = \frac{1}{2} \log_2 \frac{|\Theta_Y|}{|\Theta_Z|} \quad \text{bits/channel use} \quad (4)$$

This formula can also be found in [1, Eq. (10.98)]. Notice that this formula makes no assumption on the pairwise correlation between the signal X_i and the noise Z_j .

Next, we consider the problem from the viewpoint of communications theory and assume that the additive complex Gaussian noise \underline{Z} is independent of the signal X . For the Gaussian channel, we let $\Theta_X = \underline{S} \underline{S}^\dagger$, where $S^T = (S_1, S_2, \dots, S_n)$ is a deterministic vector with $\sum_{i=1}^n |S_i|^2 = P$, and \dagger is defined as the conjugate transpose of a vector (that is, $S^\dagger = S^{*T}$). Now the covariance matrix of the observables can be expressed as

$$\Theta_Y = \Theta_X + \Theta_Z = \underline{S} \underline{S}^\dagger + \Theta_Z \quad (5)$$

since, in this case, $ES_i Z_j^* = 0$ for $1 \leq i, j \leq n$. With the above notation, Eq. (4) can be evaluated as

$$C_{array} = \frac{1}{2} \log_2 \frac{|\underline{S} \underline{S}^\dagger + \Theta_Z|}{|\Theta_Z|} \quad \text{bits/channel use} \quad (6)$$

emphasizing the independent signal and noise components of the observable covariance matrix. Using a corollary to Theorem A.3.2 in [3], the ratio of determinants in Eq. (6) can be written as a quadratic form of the inverse noise covariance matrix and the signal vector as

$$\frac{|\underline{S} \underline{S}^\dagger + \Theta_Z|}{|\Theta_Z|} = 1 + \underline{S}^\dagger \Theta_Z^{-1} \underline{S} \quad (7)$$

and, hence, the array capacity may be equivalently expressed as

$$C_{array} = \frac{1}{2} \log_2(1 + \underline{S}^\dagger \Theta_Z^{-1} \underline{S}) \quad \text{bits/channel use} \quad (8)$$

While Eqs. (6) and (8) are mathematically equivalent, Eq. (8) is particularly important for the following reasons: First, it provides useful insights into the behavior of array capacity and, second, it suggests a simple receiver structure for processing the array observables.

IV. Receiver Structure for an Array of Gaussian Channels

Let $\underline{w}^T = (w_1, w_2, \dots, w_n)$ be a complex weight vector. It follows that the SNR of the scalar output $v = \underline{w}^T \underline{Y}$ ($\underline{Y}^T = (Y_1, Y_2, \dots, Y_n)$) is given by

$$\begin{aligned}
SNR &= \frac{|\underline{w}^T \underline{S}|^2}{E(|\underline{w}^T \underline{Z}|^2)} \\
&= \frac{\underline{w}^T \underline{S} \underline{S}^\dagger \underline{w}^*}{\underline{w} \Theta_Z \underline{w}^*}
\end{aligned} \tag{9}$$

This expression holds for any weight vector. As shown in the Appendix, the weight vector \underline{w} that maximizes the SNR is given by

$$\underline{w}_{opt} = a(\Theta_Z^{-1} \underline{S})^* \tag{10}$$

where a is an arbitrary complex constant. Substituting Eq. (10) into Eq. (9) yields the optimal SNR

$$SNR_{max} = \underline{S}^\dagger \Theta_Z^{-1} \underline{S} \tag{11}$$

which is seen to be the same as the quadratic form in Eq. (8). Thus, the quadratic form in Eq. (8) is equivalent to the maximum of the SNR obtained from an optimally weighted sum of the array observables. The array receiver structure implied by this observation is shown in Fig. 2.

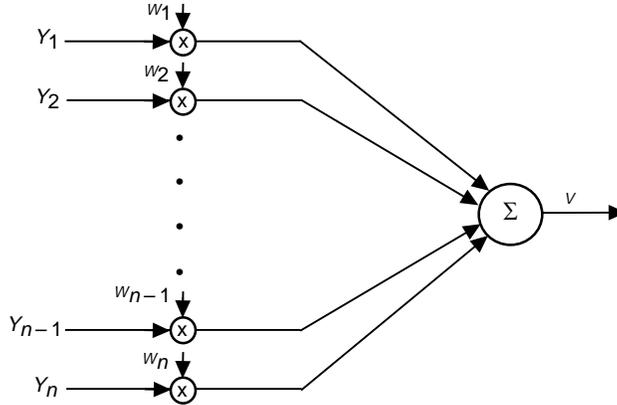


Fig. 2. Receiver structure derived from Eq. (8).

It is well known that a Gaussian source achieves capacity for a Gaussian channel [1]. Since the output of the receiver in Fig. 2 is a weighted sum of the X_i plus Gaussian noise, it is a Gaussian random variable for any value of the source X : hence, the output is a Gaussian channel. Since the array capacity in Eq. (8) depends only on the maximum SNR of the output variable v , it follows that the capacity of the scalar channel of Fig. 2 equals the capacity of the array channel of Fig. 1. This is an important observation since it enables the conversion of an n -dimensional vector observable to a single-dimensional scalar observable without decreasing the capacity of the channel.

Writing Eq. (8) in terms of Eq. (11) emphasizes the point that the channel capacity of the array depends only on the maximum SNR that can be achieved by a weighted sum of the array observables:

$$C_{array} = \frac{1}{2} \log_2(1 + SNR_{max}) \tag{12}$$

It follows that the maximum of the mutual information between v and X can also be stated directly as

$$\max_{p(x)} \{H(v) - H(v|X)\} = C_v = C_{array} \quad (13)$$

which is simply the capacity C_v of the scalar channel.

When the components of the noise vector \underline{Z} are uncorrelated, both the covariance matrix Θ_Z and its inverse Θ_Z^{-1} become diagonal matrices, with the diagonal element of the inverse matrix equal to the inverse of the corresponding diagonal element of the covariance matrix. With the i th diagonal element of Θ_Z equal to $\sigma_{Z,i}^2$, the maximized *SNR* becomes

$$SNR_{max} = \underline{S}^\dagger \Theta_Z^{-1} \underline{S} = \sum_{i=1}^n \frac{|S_i|^2}{\sigma_{Z,i}^2} \quad (14)$$

where the right-hand side of Eq. (14) is recognized as the sum of the individual channel SNRs. The array capacity follows directly as

$$C_{array} = \frac{1}{2} \log_2 \left(1 + \sum_{i=1}^n \frac{|S_i|^2}{\sigma_{Z,i}^2} \right) \quad \text{bits/channel use} \quad (15)$$

Thus, the capacity of an array of Gaussian channels with uncorrelated noise is equivalent to that of a single Gaussian channel, with an SNR equal to the sum of the individual channel SNRs.

For the correlated noise case, the observation of the noise in any channel provides useful information about the noise in every other channel. This concept is called “noise cancellation” in the adaptive signal-processing literature.

V. Examples

Some examples of simple array channels that allow closed-form analytic solutions and provide insights into the problem are considered.

A. Two-Channel Array

Consider a two-channel array with signal vector $\underline{S} = (S_1, S_2)$ and noise covariance matrix

$$\Theta_Z = \begin{bmatrix} \sigma_{Z,1}^2 & \rho \sigma_{Z,1} \sigma_{Z,2} \\ \rho^* \sigma_{Z,1} \sigma_{Z,2} & \sigma_{Z,2}^2 \end{bmatrix} \quad (16)$$

where $E Z_1 Z_2^* = \rho \sigma_{Z,1} \sigma_{Z,2}$, $E |Z_1|^2 = \sigma_{Z,1}^2$, and $E |Z_2|^2 = \sigma_{Z,2}^2$. The determinant of the noise covariance matrix is $|\Theta_Z| = \sigma_{Z,1}^2 \sigma_{Z,2}^2 (1 - |\rho|^2)$; its inverse is given by

$$\Theta_Z^{-1} = \frac{1}{1 - |\rho|^2} \begin{bmatrix} \frac{1}{\sigma_{Z,1}^2} & \frac{-\rho}{\sigma_{Z,1} \sigma_{Z,2}} \\ \frac{-\rho^*}{\sigma_{Z,1} \sigma_{Z,2}} & \frac{1}{\sigma_{Z,2}^2} \end{bmatrix} \quad (17)$$

and the resulting array channel capacity is

$$C_{array} = \frac{1}{2} \log_2 \left(1 + \frac{1}{1 - |\rho|^2} \left(\frac{|S_1|^2}{\sigma_{Z,1}^2} + \frac{|S_2|^2}{\sigma_{Z,2}^2} - \frac{2\text{Re}(\rho S_1^* S_2)}{\sigma_{Z,1} \sigma_{Z,2}} \right) \right) \quad (18)$$

As the magnitude of the correlation coefficient approaches 1, the array capacity grows without bound except for the special case when $|S_1|/\sigma_{Z,1} = |S_2|/\sigma_{Z,2}$ and the phase of ρ cancels the phase of $S_1^* S_2$. This corresponds to the singular detection case in communications theory where the signal is recovered without error in the absence of noise. If the noise is uncorrelated ($\rho = 0$), the array capacity depends only on the sum of the channel SNRs, as noted above. If the noise is correlated but the signal is 0 in one of the channels (for example, $|S_2| = 0$), the array capacity reduces to

$$C_{array} = \frac{1}{2} \log_2 \left(1 + \frac{|S_1|^2}{\sigma_{Z,1}^2(1 - |\rho|^2)} \right), \quad |S_2| = 0 \quad (19)$$

This “noise-cancellation” result is independent of the phase of the correlation coefficient, implying that the complex noise samples in the two channels need not be phase aligned for perfect cancellation—in effect, the optimum weights defined in Eq. (10) ensure that the noise phases are properly aligned. The array capacity for the two-channel noise-cancellation case corresponding to Eq. (19) is shown in Fig. 3 for several SNRs.

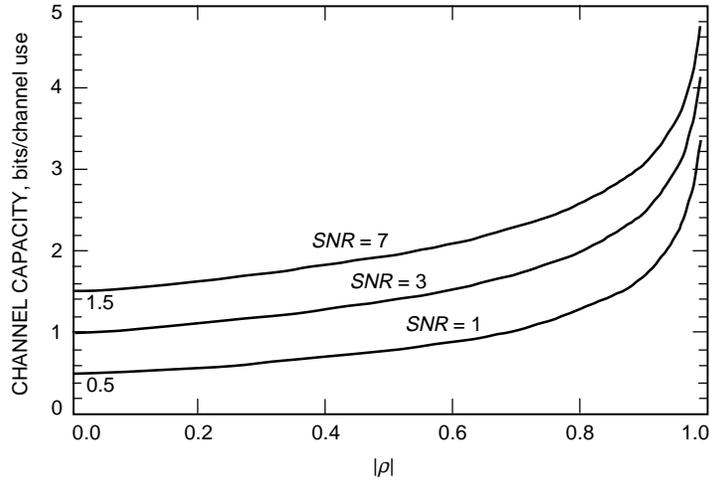


Fig. 3. Array channel for the two-channel case, $|S_2| = 0$.

B. Three-Channel Examples

Next, we consider some three-channel examples, but only for the case of real signals and noise (in other words, here we ignore phase effects). We consider the triangular and the linear constellations as shown in Fig. 4, each having an array of three elements. We assume that the correlation coefficient of Z_i and Z_j , which is denoted by ρ_{ij} , is geometrically related to the distance d_{ij} of array elements i and j as follows:

$$\rho_{ij} = \rho^{|d_{ij}|} \quad (20)$$

where $\max\{-1, a\} \leq \rho \leq \min\{1, b\}$, and a and b define the range $[a, b] \subseteq [-1, 1]$ such that the covariance matrix Θ_Z is legal, that is, Θ_Z is non-negative definite and $|\Theta_Z| \geq 0$. Also,

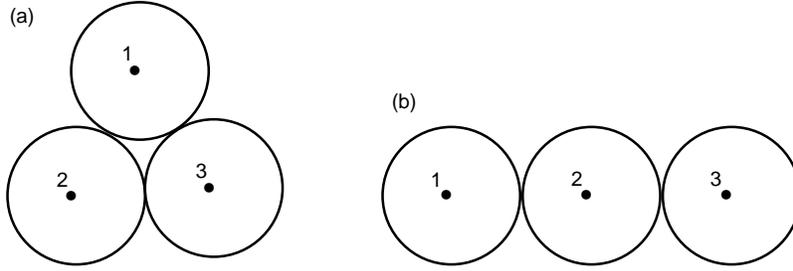


Fig. 4. Array configurations of three elements: (a) triangular configuration and (b) linear configuration.

$$EZ_i Z_j = \rho_{ij} \sigma_{Z,i} \sigma_{Z,j} \quad (21)$$

The covariance matrices Θ_Z^1 and Θ_Z^2 of the triangular and linear constellations, therefore, are given by

$$\Theta_Z^1 = \begin{bmatrix} \sigma_{Z,1}^2 & \rho \sigma_{Z,1} \sigma_{Z,2} & \rho \sigma_{Z,1} \sigma_{Z,3} \\ \rho \sigma_{Z,1} \sigma_{Z,2} & \sigma_{Z,2}^2 & \rho \sigma_{Z,2} \sigma_{Z,3} \\ \rho \sigma_{Z,1} \sigma_{Z,3} & \rho \sigma_{Z,2} \sigma_{Z,3} & \sigma_{Z,3}^2 \end{bmatrix} \quad (22)$$

and

$$\Theta_Z^2 = \begin{bmatrix} \sigma_{Z,1}^2 & \rho \sigma_{Z,1} \sigma_{Z,2} & \rho^2 \sigma_{Z,1} \sigma_{Z,3} \\ \rho \sigma_{Z,1} \sigma_{Z,2} & \sigma_{Z,2}^2 & \rho \sigma_{Z,2} \sigma_{Z,3} \\ \rho^2 \sigma_{Z,1} \sigma_{Z,3} & \rho \sigma_{Z,2} \sigma_{Z,3} & \sigma_{Z,3}^2 \end{bmatrix} \quad (23)$$

By substituting Θ_Z^1 and Θ_Z^2 into Eq. (8), we can evaluate the array channel capacities of the above array constellations as a function of ρ . We consider two cases: channels having different SNRs and channels having the same SNR.

1. Channels Having Different SNRs. Let $\sigma_{X,1}^2 = 1$, $\sigma_{X,2}^2 = 1$, and $\sigma_{X,3}^2 = 9$; let $\sigma_{Z,1}^2 = 0.16$, $\sigma_{Z,2}^2 = 0.49$, and $\sigma_{Z,3}^2 = 1.44$. The array channel capacities of the two constellations as a function of ρ are given in Fig. 5. We observe that the array channel capacities approach infinity when ρ approaches -0.5 and 1.0 for the triangular constellation (when $|\Theta_Z^1| = 0$) and when ρ approaches -1.0 and 1.0 for the linear constellation (when $|\Theta_Z^2| = 0$). The channel capacities of both constellations are the same (3.2 bits/channel use) at $\rho = 0$, and they are lower bounded at 2.75 bits/channel use for this example. Thus, the maximum degradation due to noise correlation is only 0.45 bits/channel use.

2. Channels Having the Same SNR. Let $\sigma_{X,1}^2 = 1.0$, $\sigma_{X,2}^2 = 1.0$, and $\sigma_{X,3}^2 = 0.25$; let $\sigma_{Z,1}^2 = 0.16$, $\sigma_{Z,2}^2 = 0.16$, and $\sigma_{Z,3}^2 = 0.04$. The array channel capacities of the two constellations as a function of ρ are given in Fig. 6. Again we observe that the array channel capacities approach infinity when ρ approaches -0.5 for the triangular constellation and when ρ approaches -1.0 for the linear constellation. However, both channel capacities approach 1.43 bits/channel use for the real signal and noise case, the channel capacity of a single channel, when ρ approaches 1.0. This is apparent from the fact that the receiver actually sees three scaled copies of the received signal plus noise, and this is equivalent to looking at one channel alone.

The observations described in the two examples conform to our intuition, and the general formula given in Eq. (8) predicts all these observations.

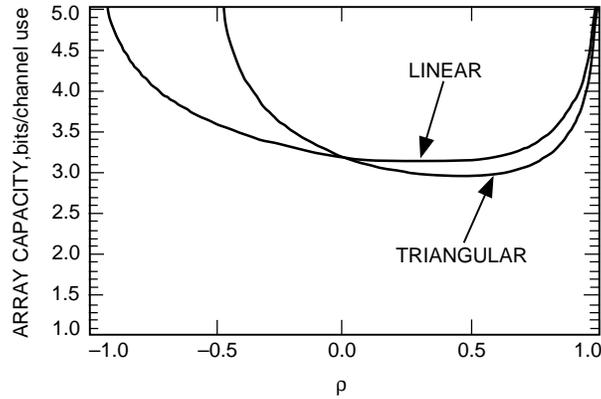


Fig. 5. Array capacity of Gaussian channels: different SNRs.

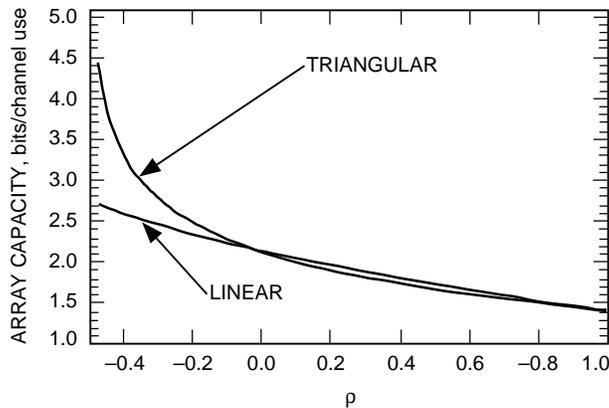


Fig. 6. Array capacity of Gaussian channels: same SNR.

VI. Summary

The capacity of an array of n Gaussian channels has been derived. The array channel was modeled as n observations of a single source, in the presence of additive Gaussian noise. It was shown that an optimally weighted sum of the array outputs achieves the same channel capacity as the array channel. Several examples of two- and three-channel arrays were discussed, and graphs of channel capacities were provided to illustrate array capacity as a function of noise correlation as well as display examples of singular behavior.

References

- [1] T. Cover and J. Thomas, *Elements of Information Theory*, New York: Wiley Interscience, 1991.
- [2] V. Vilnrotter, E. Rodemich, and S. Dolinar, "Real-Time Combining of Residual Carrier Array Signals Using ML Weight Estimates," *IEEE Trans. Comm.*, vol. 40, no. 3, pp. 604–615, March 1992.
- [3] T. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York: Wiley, 1958.

- [4] L. E. Brennan and L. S. Reed, “Theory of Adaptive Radar,” *Array Processing Applications to Radar*, ed. S. Haykin, Stroudsburg, Pennsylvania: Dowden, Hutchinson and Ross, Inc., 1980.
- [5] S. P. Applebaum, “Adaptive Arrays,” *Array Processing Applications to Radar*, ed. S. Haykin, Stroudsburg, Pennsylvania: Dowden, Hutchinson and Ross, Inc., 1980.
- [6] H. H. Tan, R. Liang, and P.-H. Suen, “Optimum Combining of Residual Carrier Array Signals in Correlated Noises,” *The Telecommunications and Data Acquisition Progress Report 42-124, October–December 1995*, Jet Propulsion Laboratory, Pasadena, California, pp. 33–52, February 15, 1996.
http://tda.jpl.nasa.gov/tda/progress_report/42-124/124K.pdf

Appendix

Derivation of Optimum Combining Weights

The following is a simple derivation of the optimum combining weights that maximize the SNR. Other derivations can be found in the adaptive signal processing literature [4,5] and in this issue [6].

We start with a derivation for the uncorrelated noise case (diagonal covariance matrix). It is shown in [2], using the Cauchy–Schwarz inequality, that for the uncorrelated case the optimum combining weight vector \underline{U} is proportional to

$$\underline{U} = (\Theta_D^{-1} \underline{V})^*$$

where \underline{V} is a signal vector, Θ_D is a diagonal matrix with components $\sigma_{Z,i}^2$, and Θ_D^{-1} is its inverse. When the weight vector \underline{U} is applied, the combined SNR is maximized, achieving its upper bound,

$$SNR_{max} = \underline{V}^\dagger \Theta_D^{-1} \underline{V}$$

Next, consider a correlated Gaussian noise vector with covariance matrix Θ_Z , and let D be a unitary matrix that diagonalizes Θ_Z . With D^\dagger the conjugate transpose of D , D^{-1} its inverse, and $D^\dagger = D^{-1}$ (unitary), we can write the diagonal covariance in terms of Θ_Z and D as

$$\Theta_D = D \Theta_Z D^T$$

Thus, D rotates vectors from the uncorrelated into the correlated reference frame, without changing their length, while its inverse rotates them in the opposite direction. Let $\underline{S} = D^{-1} \underline{V}$ be a signal vector, and let $\underline{W} = D^{-1} \underline{U}$ be a weight vector in the correlated frame. The optimum weights can be written in terms of Θ_Z and \underline{S} as

$$\begin{aligned}
\underline{U} &= (\Theta_D^{-1} \underline{V})^* = (D^* \Theta_Z^* D^\dagger)^{-1} \underline{V}^* \\
&= (\Theta_Z^* D^\dagger)^{-1} D^{*-1} \underline{V}^* \\
&= (D^\dagger)^{-1} \Theta_Z^{*-1} (D^{*-1} \underline{V}^*) \\
&= D(\Theta_Z^{-1} \underline{\mathcal{S}})^*
\end{aligned}$$

This is the optimal weight vector in the uncorrelated frame, in terms of Θ_Z^{-1} and $\underline{\mathcal{S}}$. Applying the inverse rotation operator D^{-1} , we obtain the optimum weight vector in the correlated frame as

$$\underline{W} = D^{-1} \underline{U} = D^{-1} D(\Theta_Z^{-1} \underline{\mathcal{S}})^* = (\Theta_Z^{-1} \underline{\mathcal{S}})^*$$